

JOINT LARGE DEVIATION RESULT FOR EMPIRICAL MEASURES OF THE NEAR INTERMEDIATE COLOURED RANDOM GEOMETRIC GRAPHS

Abstract. We prove joint large deviation principle for the *empirical pair measure* and *empirical locality measure* of the *near intermediate* coloured random geometric graph models, see (Canning & Penman, 2003), on n points picked uniformly in $[0, 1]^d$, for $d \in \mathbb{N}$. From this result we obtain large deviation principles for the *number of edges per vertex*, the *degree distribution* and the *proportion of isolated vertices* for the *near intermediate* random geometric graph models.

1. INTRODUCTION

In this article we study the coloured geometric random graph CGRG, where n points or vertices or nodes are picked uniformly at random in $[0, 1]^d$, colours or spins are assigned independently from a finite alphabet \mathcal{U} and any two points with colours $a_1, a_2 \in \mathcal{U}$ distance at most $r_n(a_1, a_2)$ apart are connected. This random graph models, which has the geometric random graph (see Penrose, 2003) as special case, has been suggested by see (Canning & Penman, 2003) as a possible extension to the coloured random graph studied by (OConnell, 1998), (Biggins and Penman, 2009), (Doku-Amponsah and Moerters, 2010), (Doku-Amponsah, 2006), (Bordenave and Caputo, 2013), (Mukherjee, 2013) and (Doku-Amponsah, 2014[b]).

Until recently few or no large deviation result about the CGRG have been found. Doku-Amponsah (2014[b]) proved some joint large deviation principle for empirical pair measure and the empirical locality measure of the CGRG, where n points are uniformly chosen in $[0, 1]^d$, colours or spins are assigned by drawing without replacement from the pool of $n\varpi_n(a_1)$ colours and $n\omega_n(a_1, a_2)$ edges, $a_1, a_2 \in \mathcal{U}$ are randomly inserted among the points for some colour law $\varpi_n : \mathcal{U} \rightarrow [0, 1]$ and edge law $\omega_n : \mathcal{U} \times \mathcal{U} \rightarrow [0, \infty)$.

This article presents a full large deviation principle (LDP) for the empirical pair measure and the empirical locality measure of the near intermediate RGG. Refer to (Doku-Amponsah and Moerters) for similar result for the coloured random graphs. From this large deviation results we obtain LDP for graph quantities such as *number of edges per vertex*, *the degree distribution* and *the proportion of isolated vertices* of geometric random graphs in the intermediate case. See, (OConnell, 1998), (Biggins and Penman, 2009), (Doku-Amponsah and Moerters, 2010), (Doku-Amponsah, 2006), (Bordenave and Caputo, 2013), (Mukherjee, 2013) and (Doku-Amponsah, 2014[b]) for similar result for the Erdős-Rényi graphs.

In the course of the proof of this joint LDP we obtain a joint LDP for *empirical colour measure* and *empirical pair measures* for the CGRG by the exponential change of measure techniques, see (Doku-Amponsah & Moerters, 2010) or (Doku-Amponsah, 2006). Our prove of the main result uses

AMS Subject Classification: 60F10, 05C80

Keywords: Random geometric graph, Erdős-Rényi graph, coloured random geometric graph, typed graph, joint large deviation principle, empirical pair measure, empirical measure, degree distribution, entropy, relative entropy, isolated vertices .

the method of mixture idea also used by (Doku-Amponsah & Moerters, 2010) or (Doku-Amponsah, 2006). To be specific about this idea, we mix the joint LDP of *empirical colour measure* and for *empirical pair measures*, and joint LDP for *empirical pair measure and the empirical locality measure* given the *empirical colour measure* and for *empirical pair measures*, (see Doku-Amponsah, 2015[b]) to obtain the full joint LDP for *empirical pair measure and the empirical locality measure* of CGRG models.

In the remainder of the paper we state and prove our LDP results. In Section 2 we state our LDPs, Theorem 2.2, Corollary 2.3 and Theorem 2.1. In Section 4 we combine (Doku-Amponsah, Theorem 2.1, 2014[b]) and (Doku-Amponsah, Theorem 2.1, 2014[c]) to obtain the Theorem 2.1, using the setup and result of (Biggins, 2004) to ‘mix’ the LDPs. The paper concludes with the proofs of our main results Theorem 2.2 and Corollary 2.3 which are given in Section 5.

2. STATEMENT OF THE RESULTS

2.1 The joint LDP for empirical pair measure and empirical locality measure of CGRG.

In this subsection we shall look at a more general model of random geometric graphs, the coloured RGGs in which the connectivity radius depends on the type or colour or symbol or spin of the nodes. The empirical pair measure and the empirical locality measure are our main object of study.

Given a probability measure ν on \mathcal{W} and a function $r_n: \mathcal{W} \times \mathcal{W} \rightarrow (0, 1]$ we may define the *randomly coloured random geometric graph* or simply *coloured random geometric graph* X with n vertices as follows: Pick vertices W_1, \dots, W_n at random independently according to the uniform distribution on $[0, 1]^2$. Assign to each vertex W_j colour $X(W_j)$ independently according to the *colour law* μ . Given the colours, we join any two vertices $W_i, W_j, (i \neq j)$ by an edge independently of everything else, if

$$\|W_i - W_j\| \leq r_n[X(W_i), X(W_j)].$$

In this article we shall refer to $r_n(a, b)$, for $a, b \in \mathcal{W}$ as a connection radius, and always consider

$$X = ((X(W_i), X(W_j)) : i, j = 1, 2, 3, \dots, n), E)$$

under the joint law of graph and colour. We interpret X as coloured RGG with vertices Y_1, \dots, Y_n chosen at random uniformly and independently from the vertices space $[0, 1]^2$. For the purposes of this study we restrict ourselves to the near intermediate cases i.e. the connection radius r_n satisfies the condition $nr_n^d(a, b) \rightarrow C(a, b)$ for all $a, b \in \mathcal{W}$, where $C: \mathcal{W}^2 \rightarrow [0, \infty)$ is a symmetric function, which is not identically equal to zero.

For any finite or countable set \mathcal{W} we denote by $\mathcal{P}(\mathcal{W})$ the space of probability measures, and by $\tilde{\mathcal{P}}(\mathcal{W})$ the space of finite measures on \mathcal{W} , both endowed with the weak topology. By convention we write $\mathbb{N} = \{0, 1, 2, \dots\}$.

We associate with any coloured graph X a probability measure, the *empirical colour measure* $\mathcal{L}^1 \in \mathcal{P}(\mathcal{W})$, by

$$\mathcal{L}_X^1(a) := \frac{1}{n} \sum_{j=1}^n \delta_{X(W_j)}(a), \quad \text{for } a \in \mathcal{W},$$

and a symmetric finite measure, the *empirical pair measure* $\mathcal{L}_X^2 \in \tilde{\mathcal{P}}_*(\mathcal{W}^2)$, by

$$\mathcal{L}_X^2(a, b) := \frac{1}{n} \sum_{(i, j) \in E} [\delta_{(X(W_i), X(W_j))}(a, b) + \delta_{(X(W_j), X(W_i))}(a, b)], \quad \text{for } (a, b) \in \mathcal{W}^2.$$

The total mass $\|\mathcal{L}_X^2\|$ of the empirical pair measure is $2|E|/n$. Finally we define a further probability measure, the *empirical neighbourhood measure* $\mathcal{M}_X \in \mathcal{P}(\mathcal{W} \times \mathbb{N})$, by

$$\mathcal{M}_X(a, \ell) := \frac{1}{n} \sum_{j=1}^n \delta_{(X(W_j), L(W_j))}(a, \ell), \quad \text{for } (a, \ell) \in \mathcal{W} \times \mathbb{N},$$

where $L(v) = (l^v(b), b \in \mathcal{W})$ and $l^v(b)$ is the number of vertices of colour b connected to vertex v .

For any $\mu \in \mathcal{P}(\mathcal{W} \times \mathbb{N}^{\mathcal{W}})$ we denote by μ_1 the \mathcal{W} -marginal of μ and for every $(b, a) \in \mathcal{W} \times \mathcal{W}$, let μ_2 be the law of the pair $(a, l(b))$ under the measure μ . Define the measure (finite), $\langle \mu(\cdot, \ell), l(\cdot) \rangle \in \tilde{\mathcal{P}}(\mathcal{W} \times \mathcal{W})$ by

$$\mathcal{H}_2(\mu)(b, a) := \sum_{l(b) \in \mathbb{N}} \mu_2(a, l(b)) l(b), \quad \text{for } a, b \in \mathcal{W}$$

and write $\mathcal{H}_1(\mu) = \mu_1$. We define the function $\mathcal{H}: \mathcal{P}(\mathcal{W} \times \mathbb{N}^{\mathcal{W}}) \rightarrow \mathcal{P}(\mathcal{W}) \times \tilde{\mathcal{P}}(\mathcal{W} \times \mathcal{W})$ by $\mathcal{H}(\mu) = (\mathcal{H}_1(\mu), \mathcal{H}_2(\mu))$ and note that $\mathcal{H}(\mathcal{M}_X) = (\mathcal{L}_X^1, \mathcal{L}_X^2)$. Observe that \mathcal{H}_1 is a continuous function but \mathcal{H}_2 is *discontinuous* in the weak topology. In particular, in the summation $\sum_{l(b) \in \mathbb{N}} \mu_2(a, l(b)) l(b)$ the

function $l(b)$ may be unbounded and so the functional $\mu \rightarrow \mathcal{H}_2(\mu)$ would not be continuous in the weak topology. We call a pair of measures $(\varpi, \mu) \in \tilde{\mathcal{P}}(\mathcal{W} \times \mathcal{W}) \times \mathcal{P}(\mathcal{W} \times \mathbb{N}^{\mathcal{W}})$ *sub-consistent* if

$$\mathcal{H}_2(\mu)(b, a) \leq \varpi(b, a), \quad \text{for all } a, b \in \mathcal{W}, \quad (2.1)$$

and *consistent* if equality holds in (2.1). For a measure $\varpi \in \tilde{\mathcal{P}}_*(\mathcal{W}^2)$ and a measure $\omega \in \mathcal{P}(\mathcal{W})$, define

$$\mathfrak{H}_C^d(\varpi \parallel \omega) := H(\varpi \parallel \rho(d)C\omega \otimes \omega) + \rho(d)\|C\omega \otimes \omega\| - \|\varpi\|,$$

where the measure $C\omega \otimes \omega \in \tilde{\mathcal{P}}(\mathcal{W} \times \mathcal{W})$ is defined by $C\omega \otimes \omega(a, b) = C(a, b)\omega(a)\omega(b)$ for $a, b \in \mathcal{W}$. It is not hard to see that $\mathfrak{H}_C^d(\varpi \parallel \omega) \geq 0$ and equality holds if and only if $\varpi = \rho(d)C\omega \otimes \omega$. For every $(\varpi, \mu) \in \tilde{\mathcal{P}}_*(\mathcal{W} \times \mathcal{W}) \times \mathcal{P}(\mathcal{W} \times \mathbb{N})$ define a probability measure $Q = Q[\varpi, \mu]$ on $\mathcal{W} \times \mathbb{N}$ by

$$Q(a, \ell) := \mu_1(a) \prod_{b \in \mathcal{W}} e^{-\frac{\varpi(a, b)}{\mu_1(a)}} \frac{1}{\ell(b)!} \left(\frac{\varpi(a, b)}{\mu_1(a)} \right)^{\ell(b)}, \quad \text{for } a \in \mathcal{W}, \ell \in \mathbb{N}.$$

We assume $d \geq 2$ is finite and write

$$\rho(d) = \frac{\pi^{d/2}}{\Gamma\left(\frac{(d+2)}{2}\right)},$$

where *Gamma* is the gamma function.

We now state the principal theorem in this section the LDP for the empirical pair measure and the empirical neighbourhood measure.

Theorem 2.1. *Suppose that X is a coloured RGG graph with colour law μ and connection radii $r_n: \mathcal{W} \times \mathcal{W} \rightarrow [0, 1]$ satisfying $nr_n^d(a, b) \rightarrow C(a, b)$ for some symmetric function $C: \mathcal{W} \times \mathcal{W} \rightarrow [0, \infty)$ not identical to zero. Then, as $n \rightarrow \infty$, the pair $(\mathcal{L}_X^2, \mathcal{M}_X)$ satisfies an LDP in $\tilde{\mathcal{P}}_*(\mathcal{W} \times \mathcal{W}) \times \mathcal{P}(\mathcal{W} \times \mathbb{N})$ with good rate function*

$$J(\varpi, \mu) = \begin{cases} H(\mu \parallel Q) + H(\mu_1 \parallel \nu) + \frac{1}{2} \mathfrak{H}_C^d(\varpi \parallel \mu_1) & \text{if } (\varpi, \mu) \text{ consistent and } \mu_1 = \varpi_2, \\ \infty & \text{otherwise.} \end{cases}$$

Remark 1 Note that on typical coloured RGG graph we have, $\omega = \mu_1$, $\varpi = \rho(d)C\mu \otimes \mu$ and

$$\mu(a, \ell) = \nu(a) \prod_{b \in \mathcal{W}} e^{-\rho(d)C(a, b)\nu(b)} \frac{(\rho(d)C(a, b)\nu(b))^{\ell(b)}}{\ell(b)!}, \quad \text{for all } (a, \ell) \in \mathcal{W} \times \mathbb{N}.$$

This is the law of a pair (a, ℓ) where a is distributed according to μ and, given the value of a , the random variables $\ell(b)$ are independently Poisson distributed with parameter $\rho(d)C(a, b)\nu(b)$. Hence, as $n \rightarrow \infty$, the empirical neighbourhood measure $\mathcal{M}_X(a, \ell)$ converges to $\mu(a, \ell)$ in probability.

Corollary 2.2. *Suppose D is the degree distribution of the random graph $\mathcal{G}(n, r_n)$, where the connectivity radius $r_n \in (0, 1]$ satisfies $nr_n^d \rightarrow c \in (0, \infty)$. Then, as $n \rightarrow \infty$, D satisfies an LDP in the space $\mathcal{P}(\mathbb{N} \cup \{0\})$ with good rate function*

$$\eta_1(\delta) = \begin{cases} \frac{1}{2} \langle \delta \rangle \log \left(\frac{\langle \delta \rangle}{\rho(d)c} \right) - \frac{1}{2} \langle \delta \rangle + \frac{\rho(d)c}{2} + H(d \| q_{\langle \delta \rangle}), & \text{if } \langle \delta \rangle < \infty, \\ \infty & \text{if } \langle \delta \rangle = \infty, \end{cases} \quad (2.2)$$

where q_x is a Poisson distribution with parameter x and $\langle \delta \rangle := \sum_{m=0}^{\infty} m\delta(m)$.

Next we give a similar result as in O'Connell [], the LDP for the proportion of isolated vertices of the RGG.

Corollary 2.3. *Suppose D is the degree distribution of the random graph $\mathcal{G}(n, r_n)$, where the connectivity radius $r_n \in (0, 1]$ satisfies $nr_n^d \rightarrow c \in (0, \infty)$. Then, as $n \rightarrow \infty$, the proportion of isolated vertices, $D(0)$ satisfies an LDP in $[0, 1]$ with good rate function*

$$\xi_1(y) = y \log y + \rho(d)cy(1 - y/2) - (1 - y) \left[\log \left(\frac{\rho(d)c}{a} \right) - \frac{(a - \rho(d)c(1 - y))^2}{2\rho(d)c(1 - y)} \right],$$

where $a = a(y)$ is the unique positive solution of $1 - e^{-a} = \frac{\rho(d)c}{a}(1 - y)$.

From Lemma 2.3 we deduce that on a typical random geometric graphs the number of isolated vertices will grow like $ne^{-\rho(d)c}$. Thus, as $n \rightarrow \infty$, the number of isolated vertices in the R.G graphs converges to $ne^{-\rho(d)c}$ in probability. In our last theorem in this subsection we give the LDP for the proportion of edges to the number of vertices of the R.G.

2.2 The joint LDP for the empirical colour measure and empirical pair measure of CGRG

Theorem 2.4. *Suppose that X is a coloured RGG graph with colour law ν and connection radii $r_n: \mathcal{Y}^2 \rightarrow [0, 1]$ satisfying $nr_n^d(a, b) \rightarrow C(a, b)$ for some symmetric function $C: \mathcal{Y}^2 \rightarrow [0, \infty)$ not identical to zero. Then, as $n \rightarrow \infty$, the pair $(\mathcal{L}_X^1, \mathcal{L}_X^2)$ satisfies an LDP in $\mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2)$ with good rate function*

$$I(\varpi, \varpi) = H(\varpi \| \nu) + \frac{1}{2} \mathfrak{H}_C^d(\varpi \| \varpi), \quad (2.3)$$

where

$$\mathfrak{H}_C^d(\varpi \| \varpi) := H(\varpi \| \Delta(d)C\varpi \otimes \varpi) + \Delta(d)\|C\varpi \otimes \varpi\| - \|\varpi\|,$$

and the measure $C\varpi \otimes \varpi \in \tilde{\mathcal{W}}(\mathcal{Y} \times \mathcal{Y})$ is defined by $C\varpi \otimes \varpi(a, b) = C(a, b)\varpi(a)\varpi(b)$ for $a, b \in \mathcal{Y}$.

Further, we state a Corollary of Theorem 2.4 below.

Corollary 2.5. *Suppose that X is a coloured RGG graph with colour law ν and connection radii $r_n: \mathcal{Y}^2 \rightarrow [0, 1]$ satisfying $nr_n^d(a, b) \rightarrow C(a, b)$ for some symmetric function $C: \mathcal{Y}^2 \rightarrow [0, \infty)$ not identical to zero. Then, as $n \rightarrow \infty$, the number of edges per vertex $|E|/n$ of X satisfies an LDP in $[0, \infty)$ with good rate function*

$$\zeta(x) = x \log x - x + \inf_{y > 0} \{ \psi(y) - x \log(y) + y \},$$

where $\psi(y) = \inf H(\varpi \| \nu)$ over all probability vectors ϖ with $\frac{1}{2}\Delta(d)\varpi^T C\varpi = y$.

Remark 2 By taking $C(a, b) = c$ one will obtain $\psi(y) = 0$ for $y = \frac{\Delta(d)}{2}c$, and $\psi(y) = \infty$ otherwise, which establishes that $|E|/n$ obeys an LDP in $[0, \infty)$ with good rate function

$$\zeta(x) = x \log x - x + \inf_{y>0} \left\{ \psi(y) - x \log\left(\frac{1}{2}y\right) + \frac{1}{2}y \right\},$$

where $\Delta(d)c = y$.

3. PROOF OF THEOREM 2.4

3.1 Change-of-Measure

For any two points U_1 and U_2 uniformly and independently chosen from the space $[0, 1]^d$ write

$$F(t) := \mathbb{P}\left\{\|U_1 - U_2\| \leq t\right\}.$$

Further, given a function $\tilde{f}: \mathcal{Y} \rightarrow \mathbb{R}$ and a symmetric function $\tilde{g}: \mathcal{Y}^2 \rightarrow \mathbb{R}$, we define the constant $U_{\tilde{f}}$ by

$$U_{\tilde{f}} = \log \sum_{a \in \mathcal{Y}} e^{\tilde{f}(a)} \nu(a),$$

and the function $\tilde{h}_n: \mathcal{Y}^2 \rightarrow \mathbb{R}$ by

$$\tilde{h}_n(a, b) = \log \left[(1 - F(r_n(a, b)) + F(r_n(a, b))e^{\tilde{g}(a, b)})^{-n} \right], \quad (3.1)$$

for $a, b \in \mathcal{Y}$. We use \tilde{f} and \tilde{g} to define (for sufficiently large n) a new coloured random graph as follows:

- To the n labelled vertices in V we assign colours from \mathcal{Y} independently and identically according to the colour law $\tilde{\nu}$ defined by

$$\tilde{\nu}(a) = e^{\tilde{f}(a) - B_{\tilde{f}}} \nu(a).$$

- Given any two vertices $u, v \in V$, with u carrying colour a and v carrying colour b , we connect vertex u to vertex v with probability

$$F(\tilde{r}_n(a, b)) = \frac{F(r_n(a, b))e^{\tilde{g}(a, b)}}{1 - F(r_n(a, b)) + F(r_n(a, b))e^{\tilde{g}(a, b)}}.$$

We denote the transformed law by $\tilde{\mathbb{P}}$. We observe that $\tilde{\nu}$ is a probability measure and that $\tilde{\mathbb{P}}$ is absolutely continuous with respect to \mathbb{P} as, for any coloured graph $X = ((X(v): v \in V), E)$,

$$\begin{aligned} \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}(X) &= \prod_{u \in V} \frac{\tilde{\nu}(X(Y_u))}{\nu(X(Y_u))} \prod_{(u, v) \in E} \frac{F(\tilde{r}_n(X(Y_u), X(Y_v)))}{F(r_n(X(Y_u), X(Y_v)))} \prod_{(u, v) \notin E} \frac{1 - F(\tilde{r}_n(X(Y_u), X(Y_v)))}{1 - F(r_n(X(Y_u), X(Y_v)))} \\ &= \prod_{u \in V} \frac{\tilde{\nu}(X(Y_u))}{\nu(X(Y_u))} \prod_{(u, v) \in E} \frac{F(\tilde{r}_n(X(Y_u), X(Y_v)))}{F(r_n(X(Y_u), X(Y_v)))} \times \frac{n - nF(r_n(X(Y_u), X(Y_v)))}{n - nF(\tilde{r}_n(X(Y_u), X(Y_v)))} \prod_{(u, v) \in \mathcal{E}} \frac{n - nF(\tilde{r}_n(X(Y_u), X(Y_v)))}{n - nF(r_n(X(Y_u), X(Y_v)))} \\ &= \prod_{u \in V} e^{\tilde{f}(X(Y_u)) - U_{\tilde{f}}} \prod_{(u, v) \in E} e^{\tilde{g}(X(Y_u), X(Y_v))} \prod_{(u, v) \in \mathcal{E}} e^{\frac{1}{n} \tilde{h}_n(X(Y_u), X(Y_v))} \\ &= \exp \left(n \langle \mathcal{L}_X^1, \tilde{f} - U_{\tilde{f}} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^2, \tilde{g} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^1 \otimes \mathcal{L}_X^1, \tilde{h}_n \rangle - \langle \frac{1}{2} L_\Delta^1, \tilde{h}_n \rangle \right), \end{aligned} \quad (3.2)$$

where

$$L_\Delta^1 = \frac{1}{n} \sum_{u \in V} \delta_{(X(Y_u), X(Y_u))}.$$

We write $\langle g, \varpi \rangle := \sum_{a,b \in \mathcal{Y}} g(a,b) \varpi(a,b)$ for $\varpi \in \tilde{\mathcal{W}}(\mathcal{Y}^2)$, and $\langle f, \varpi \rangle := \sum_{a \in \mathcal{Y}} f(a) \varpi(a)$ for $\varpi \in \mathcal{W}(\mathcal{Y})$, and note that

$$F(r_n(a,b)) = \Delta(d)r_n^d(a,b), \text{ for all } a,b \in \mathcal{Y}^2.$$

i.e. the volume of a d -dimensional (hyper)sphere with radius $r(a,b)$ satisfying $nr_n^d(a,b) \rightarrow C(a,b)$.

The following lemmas will be useful in the proofs of main Lemmas.

Lemma 3.1 (Euler's lemma). *If $nr_n^d(a,b) \rightarrow C(a,b)$ for every $a,b \in \mathcal{Y}$, then*

$$\lim_{n \rightarrow \infty} [1 + \alpha F(r_n(a,b))]^n = e^{\alpha \Delta(d)C(a,b)}, \text{ for all } a,b \in \mathcal{Y} \text{ and } \alpha \in \mathbb{R}. \quad (3.3)$$

Proof. Observe that, for any $\varepsilon > 0$ and for large n we have

$$\left[1 + \frac{\alpha \Delta(d)C(a,b) - \varepsilon}{n}\right]^n \leq [1 + \alpha F(r_n(a,b))]^n \leq \left[1 + \frac{\alpha \Delta(d)C(a,b) + \varepsilon}{n}\right]^n,$$

by the pointwise convergence. Hence by the sandwich theorem and Euler's formula we get (3.3). ■

We write

$$P^{(n)}(\varpi) := \mathbb{P}\{\mathcal{L}_X^1 = \varpi\}.$$

Lemma 3.2. *The family of measures $(P^n: n \in \mathbb{N})$ is exponentially tight on $\mathcal{W}(\mathcal{Y})$*

Proof. We use coupling argument, see the proof of [, Lemma 5.1] to show that, for every $\theta > 0$, there exists $N \in \mathbb{N}$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{P}\{|E| > nN\} \leq -\theta.$$

To begin, let $c > \max_{a,b \in \mathcal{Y}} C(a,b) > 0$. Using similar coupling arguments as in see the proof of [, Lemma 5.1], we can define, for all sufficiently large n , a new coloured geometric random graph \tilde{X} with vertices Y_1, Y_2, \dots, Y_n chosen uniformly from the vertices space $[0,1]^d$, colour law μ and connectivity radius $\left(\frac{c}{n}\right)^{1/d}$, such that any edge present in X is also present in \tilde{X} . Let $|\tilde{E}|$ be the number of edges of \tilde{X} . Using the binomial formula and Euler's formula, we have that

$$\begin{aligned} \mathbb{P}\{|\tilde{E}| \geq nl\} &\leq e^{-nl} \mathbb{E}[e^{|\tilde{E}|}] = e^{-nl} \sum_{k=0}^{\frac{n(n-1)}{2}} e^k \binom{n(n-1)/2}{k} \left(\frac{\Delta(d)c}{n}\right)^k \left(1 - \frac{\Delta(d)c}{n}\right)^{n(n-1)/2-k} \\ &= e^{-nl} \left(1 - \frac{c\Delta(d)}{n} + e \frac{c\Delta(d)}{n}\right)^{n(n-1)/2} \leq e^{-nl} e^{nc\Delta(d)(e-1+o(1))}. \end{aligned}$$

Now given $\theta > 0$ choose $N \in \mathbb{N}$ such that $N > \theta + \Delta(d)c(e-1)$ and observe that, for sufficiently large n ,

$$\mathbb{P}\{|E| \geq nN\} \leq \mathbb{P}\{|\tilde{E}| \geq nN\} \leq e^{-n\theta},$$

which implies the statement. ■

3.2 Proof of the upper bound in Theorem 2.4

We denote by \mathcal{C}_1 the space of functions on \mathcal{Y} and by \mathcal{C}_2 the space of symmetric functions on \mathcal{Y}^2 , and define

$$\hat{I}(\varpi, \varpi) = \sup_{\substack{f \in \mathcal{C}_1 \\ g \in \mathcal{C}_2}} \left\{ \sum_{a \in \mathcal{Y}} (f(a) - U_f) \varpi(a) + \frac{1}{2} \sum_{a,b \in \mathcal{Y}} g(a,b) \varpi(a,b) + \frac{\Delta(d)}{2} \sum_{a,b \in \mathcal{Y}} (1 - e^{g(a,b)}) C(a,b) \varpi(a) \varpi(b) \right\}$$

for $(\varpi, \varpi) \in \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2)$.

Lemma 3.3. *For each closed set $G \subset \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2)$, we have*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in F\} \leq - \inf_{(\varpi, \varpi) \in F} \hat{I}(\varpi, \varpi).$$

Proof. First let $\tilde{f} \in \mathcal{C}_1$ and $\tilde{g} \in \mathcal{C}_2$ be arbitrary. Define $\tilde{\beta}: \mathcal{Y}^2 \rightarrow \mathbb{R}$ by

$$\tilde{\beta}(a, b) = \Delta(d)(1 - e^{\tilde{g}(a, b)})C(a, b).$$

Observe that, by Lemma 3.1, $\tilde{\beta}(a, b) = \lim_{n \rightarrow \infty} \tilde{h}_n(a, b)$ for all $a, b \in \mathcal{Y}$, recalling the definition of \tilde{h}_n from (3.1). Hence, by (3.2), for sufficiently large n ,

$$e^{\max_{a \in \mathcal{Y}} |\tilde{\beta}(a, a)|} \geq \int e^{\langle \frac{1}{2} L_\Delta^1, \tilde{h}_n \rangle} d\tilde{\mathbb{P}} = \mathbb{E} \left\{ e^{n \langle \mathcal{L}_X^1, \tilde{f} - U_{\tilde{f}} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^2, \tilde{g} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^1 \otimes \mathcal{L}_X^1, \tilde{h}_n \rangle} \right\},$$

where $L_\Delta^1 = \frac{1}{n} \sum_{u \in V} \delta_{(X(Y_u), X(Y_u))}$ and therefore,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \left\{ e^{n \langle \mathcal{L}_X^1, \tilde{f} - U_{\tilde{f}} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^2, \tilde{g} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^1 \otimes \mathcal{L}_X^1, \tilde{h}_n \rangle} \right\} \leq 0. \quad (3.4)$$

Given $\varepsilon > 0$ let $\hat{I}_\varepsilon(\varpi, \varpi) = \min\{\hat{I}(\varpi, \varpi), \varepsilon^{-1}\} - \varepsilon$. Suppose that $(\varpi, \varpi) \in G$ and observe that $\hat{I}(\varpi, \varpi) > \hat{I}_\varepsilon(\varpi, \varpi)$. We now fix $\tilde{f} \in \mathcal{C}_1$ and $\tilde{g} \in \mathcal{C}_2$ such that

$$\langle \tilde{f} - U_{\tilde{f}}, \varpi \rangle + \frac{1}{2} \langle \tilde{g}, \varpi \rangle + \frac{1}{2} \langle \tilde{\beta}, \varpi \otimes \varpi \rangle \geq \hat{I}_\varepsilon(\varpi, \varpi).$$

As \mathcal{Y} is finite, there exist open neighbourhoods B_ϖ^2 and B_ϖ^1 of ϖ, ϖ such that

$$\inf_{\substack{\tilde{\omega} \in B_\varpi^1 \\ \tilde{\omega} \in B_\varpi^2}} \left\{ \langle \tilde{f} - U_{\tilde{f}}, \tilde{\omega} \rangle + \frac{1}{2} \langle \tilde{g}, \tilde{\omega} \rangle + \frac{1}{2} \langle \tilde{\beta}, \tilde{\omega} \otimes \tilde{\omega} \rangle \right\} \geq \hat{I}_\varepsilon(\varpi, \varpi) - \varepsilon.$$

Using Chebyshev's inequality and (3.4) we have that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \left\{ e^{n \langle \mathcal{L}_X^1, \tilde{f} - U_{\tilde{f}} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^2, \tilde{g} \rangle + n \langle \frac{1}{2} \mathcal{L}_X^1 \otimes \mathcal{L}_X^1, \tilde{h}_n \rangle} \right\} - \hat{I}_\varepsilon(\varpi, \varpi) + \varepsilon \\ & \leq -\hat{I}_\varepsilon(\varpi, \varpi) + \varepsilon. \end{aligned} \quad (3.5)$$

Now we use Lemma 3.2 with $\theta = \varepsilon^{-1}$, to choose $N(\varepsilon) \in \mathbb{N}$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{|E| > nN(\varepsilon)\} \leq -\varepsilon^{-1}. \quad (3.6)$$

For this $N(\varepsilon)$, define the set $K_{N(\varepsilon)}$ by

$$K_{N(\varepsilon)} = \left\{ (\varpi, \varpi) \in \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2) : \|\varpi\| \leq 2N(\varepsilon) \right\},$$

and recall that $\|\mathcal{L}_X^2\| = 2|E|/n$. The set $K_{N(\varepsilon)} \cap F$ is compact and therefore may be covered by finitely many sets $B_{\varpi_r}^1 \times B_{\varpi_r}^2, r = 1, \dots, m$ with $(\varpi_r, \varpi_r) \in F$ for $r = 1, \dots, m$. Consequently,

$$\mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in F\} \leq \sum_{r=1}^m \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_{\varpi_r}^1 \times B_{\varpi_r}^2\} + \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \notin K_{N(\varepsilon)}\}.$$

We may now use (3.5) and (3.6) to obtain, for all sufficiently small $\varepsilon > 0$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in F\} &\leq \max_{r=1}^m \left(\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_{\varpi_r}^1 \times B_{\varpi_r}^2\} \right) \vee (-\varepsilon)^{-1} \\ &\leq \left(- \inf_{(\varpi, \varpi) \in G} \hat{I}_\varepsilon(\varpi, \varpi) + \varepsilon \right) \vee (-\varepsilon)^{-1}. \end{aligned}$$

Taking $\varepsilon \downarrow 0$ we get the desired statement. \blacksquare

Next, we express the rate function in term of relative entropies, see for example [, (2.15)], and consequently show that it is a good rate function. Recall the definition of the function I from Theorem 2.4.

Lemma 3.4.

- (i) $\hat{I}(\varpi, \varpi) = I(\varpi, \varpi)$, for any $(\varpi, \varpi) \in \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2)$,
- (ii) I is a good rate function and
- (iii) $\mathfrak{H}_C^d(\varpi \parallel \varpi) \geq 0$ with equality if and only if $\varpi = \Delta(d)C\varpi \otimes \varpi$.

Proof. (i) Suppose that $\varpi \not\ll \Delta(d)C\varpi \otimes \varpi$. Then, there exists $a_0, b_0 \in \mathcal{Y}$ with $C\varpi \otimes \varpi(a_0, b_0) = 0$ and $\varpi(a_0, b_0) > 0$. Define $\hat{g}: \mathcal{Y}^2 \rightarrow \mathbb{R}$ by

$$\hat{g}(a, b) = \log [K(\mathbb{1}_{(a_0, b_0)}(a, b) + \mathbb{1}_{(b_0, a_0)}(a, b)) + 1], \text{ for } a, b \in \mathcal{Y} \text{ and } K > 0.$$

For this choice of \hat{g} and $f = 0$ we have

$$\begin{aligned} &\sum_{a \in \mathcal{Y}} (f(a) - U_f) \varpi(a) + \sum_{a, b \in \mathcal{Y}} \frac{1}{2} \hat{g}(a, b) \varpi(a, b) + \sum_{a, b \in \mathcal{Y}} \frac{\Delta(d)}{2} (1 - e^{\hat{g}(a, b)}) C(a, b) \varpi(a) \varpi(b) \\ &\geq \frac{\Delta(d)}{2} \log(K + 1) \varpi(a_0, b_0) \rightarrow \infty, \quad \text{for } K \uparrow \infty. \end{aligned}$$

Now suppose that $\varpi \ll C\varpi \otimes \varpi$. We have

$$\begin{aligned} \hat{I}(\varpi, \varpi) &= \sup_{f \in \mathcal{C}_1} \left\{ \sum_{a \in \mathcal{Y}} \left(f(a) - \log \sum_{a \in \mathcal{Y}} e^{f(a)} \nu(a) \right) \varpi(a) \right\} \\ &\quad + \frac{\Delta(d)}{2} \sum_{a, b \in \mathcal{Y}} C(a, b) \varpi(a) \varpi(b) + \frac{1}{2} \sup_{g \in \mathcal{C}_2} \left\{ \sum_{a, b \in \mathcal{Y}} g(a, b) \varpi(a, b) - \Delta(d) \sum_{a, b \in \mathcal{Y}} e^{g(a, b)} C(a, b) \varpi(a) \varpi(b) \right\}. \end{aligned}$$

By the variational characterization of relative entropy, the first term equals $H(\varpi \parallel \nu)$. By the substitution $h = \Delta(d)e^g \frac{C\varpi \otimes \varpi}{\varpi}$ the last term equals

$$\begin{aligned} &\sup_{\substack{h \in \mathcal{C}_2 \\ h \geq 0}} \sum_{a, b \in \mathcal{Y}} \left[\log \left(h(a, b) \frac{\varpi(a, b)}{\Delta(d)C(a, b)\varpi(a)\varpi(b)} \right) - h(a, b) \right] \varpi(a, b) \\ &= \sup_{\substack{h \in \mathcal{C}_2 \\ h \geq 0}} \sum_{a, b \in \mathcal{Y}} (\log h(a, b) - h(a, b)) \varpi(a, b) + \sum_{a, b \in \mathcal{Y}} \log \left(\frac{\varpi(a, b)}{\Delta(d)C(a, b)\varpi(a)\varpi(b)} \right) \varpi(a, b) \\ &= -\|\varpi\| + H(\varpi \parallel \Delta(d)C\varpi \otimes \varpi), \end{aligned}$$

where we have used $\sup_{x>0} \log x - x = -1$ in the last step. This yields that $\hat{I}(\varpi, \varpi) = I(\varpi, \varpi)$.

(ii) Recall from (2.3) and the definition of \mathfrak{H}_C that $I(\varpi, \varpi) = H(\varpi \parallel \nu) + \frac{1}{2} H(\varpi \parallel \Delta(d)C\varpi \otimes \varpi) + \frac{\Delta(d)}{2} \|C\varpi \otimes \varpi\| - \frac{1}{2} \|\varpi\|$. All summands are continuous in ϖ, ϖ and thus I is a rate function. Moreover, for all $\alpha < \infty$, the level sets $\{I \leq \alpha\}$ are contained in the bounded set $\{(\varpi, \varpi) \in \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2): \mathfrak{H}_C^d(\varpi \parallel \varpi) \leq \alpha\}$ and are therefore compact. Consequently, I is a good rate function.

(iii) Consider the nonnegative function $\xi(x) = x \log x - x + 1$, for $x > 0$, $\xi(0) = 1$, which has its only root in $x = 1$. Note that

$$\mathfrak{H}_C^d(\varpi \parallel \varpi) = \begin{cases} \int \xi \circ g \, d(\Delta(d)C\varpi \otimes \varpi) & \text{if } g := \frac{d\varpi}{(\Delta(d)C\varpi \otimes \varpi)} \geq 0 \text{ exists,} \\ \infty & \text{otherwise.} \end{cases} \quad (3.7)$$

Hence $\mathfrak{H}_C^d(\varpi \parallel \varpi) \geq 0$, and, if $\varpi = \Delta(d)C\varpi \otimes \varpi$, then $\xi(\frac{d\varpi}{(\Delta(d)C\varpi \otimes \varpi)}) = \xi(1) = 0$ and so $\mathfrak{H}_C^d(\Delta(d)C\varpi \otimes \varpi \parallel \varpi) = 0$. Conversely, if $\mathfrak{H}_C^d(\varpi \parallel \varpi) = 0$, then $\varpi(a, b) > 0$ implies $C\varpi \otimes \varpi(a, b) > 0$, which then implies $\xi \circ g(a, b) = 0$ and further $g(a, b) = 1$. Hence $\varpi = \Delta(d)C\varpi \otimes \varpi$, which completes the proof of (iii). \blacksquare

3.3 Proof of the lower bound in Theorem 2.4

We obtain the lower bound of Theorem 2.4 from the upper bound as follows:

Lemma 3.5. *For every open set $O \subset \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2)$, we have*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in O \right\} \geq - \inf_{(\varpi, \varpi) \in O} I(\varpi, \varpi).$$

Proof. Suppose $(\varpi, \varpi) \in O$, with $\varpi \ll \Delta(d)C\varpi \otimes \varpi$. Define $\tilde{f}_\varpi: \mathcal{Y} \rightarrow \mathbb{R}$ by

$$\tilde{f}_\varpi(a) = \begin{cases} \log \frac{\varpi(a)}{\nu(a)}, & \text{if } \varpi(a) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

and $\tilde{g}_\varpi: \mathcal{Y}^2 \rightarrow \mathbb{R}$ by

$$\tilde{g}_\varpi(a, b) = \begin{cases} \log \frac{\varpi(a, b)}{\Delta(d)C(a, b)\varpi(a)\varpi(b)}, & \text{if } \varpi(a, b) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

In addition, we let $\tilde{\beta}_\varpi(a, b) = \Delta(d)C(a, b)(1 - e^{\tilde{g}_\varpi(a, b)})$ and note that $\tilde{\beta}_\varpi(a, b) = \lim_{n \rightarrow \infty} \tilde{h}_{\varpi, n}(a, b)$, for all $a, b \in \mathcal{Y}$ where

$$\tilde{h}_{\varpi, n}(a, b) = \log \left[(1 - F(r_n(a, b)) + F(r_n(a, b))e^{\tilde{g}_\varpi(a, b)})^{-n} \right].$$

Choose B_ϖ^1, B_ϖ^2 open neighbourhoods of ϖ, ϖ , such that $B_\varpi^1 \times B_\varpi^2 \subset O$ and for all $(\tilde{\varpi}, \tilde{\varpi}) \in B_\varpi^1 \times B_\varpi^2$

$$\langle \tilde{f}_\varpi, \varpi \rangle + \frac{1}{2} \langle \tilde{g}_\varpi, \varpi \rangle + \frac{1}{2} \langle \tilde{\beta}_\varpi, \varpi \otimes \varpi \rangle - \varepsilon \leq \langle \tilde{f}_{\tilde{\varpi}}, \tilde{\varpi} \rangle + \frac{1}{2} \langle \tilde{g}_{\tilde{\varpi}}, \tilde{\varpi} \rangle + \frac{1}{2} \langle \tilde{\beta}_{\tilde{\varpi}}, \tilde{\varpi} \otimes \tilde{\varpi} \rangle.$$

We now use $\tilde{\mathbb{P}}$, the probability measure obtained by transforming \mathbb{P} using the functions $\tilde{f}_\varpi, \tilde{g}_\varpi$. Note that the colour law in the transformed measure is now ϖ , and the connectivity radii $\tilde{r}_n(a, b)$ satisfy

$$n \tilde{r}_n^d(a, b) \rightarrow \varpi(a, b) / (\varpi(a)\varpi(b)) =: \tilde{C}(a, b), \text{ as } n \rightarrow \infty.$$

Using (3.2), we obtain

$$\begin{aligned} \mathbb{P} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in O \right\} &\geq \tilde{\mathbb{E}} \left\{ \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}(X) \mathbb{1}_{\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2\}} \right\} \\ &= \tilde{\mathbb{E}} \left\{ \prod_{u \in V} e^{-\tilde{f}_\varpi(X(Y_u))} \prod_{(u, v) \in E} e^{-\tilde{g}_\varpi(X(Y_u), X(Y_v))} \prod_{(u, v) \in \mathcal{E}} e^{-\frac{1}{n} \tilde{h}_{\varpi, n}(X(Y_u), X(Y_v))} \mathbb{1}_{\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2\}} \right\} \\ &= \tilde{\mathbb{E}} \left\{ e^{-n \langle \mathcal{L}_X^1, \tilde{f}_\varpi \rangle - n \frac{1}{2} \langle \mathcal{L}_X^2, \tilde{g}_\varpi \rangle - n \frac{1}{2} \langle \mathcal{L}_X^1 \otimes \mathcal{L}_X^1, \tilde{\beta}_\varpi \rangle + \frac{1}{2} \langle L_\Delta^1, \tilde{h}_{\varpi, n} \rangle} \times \mathbb{1}_{\{(\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2\}} \right\} \\ &\geq \exp \left(-n \langle \tilde{f}_\varpi, \varpi \rangle - n \frac{1}{2} \langle \tilde{g}_\varpi, \varpi \rangle - n \frac{1}{2} \langle \tilde{\beta}_\varpi, \varpi \otimes \varpi \rangle + m - n\varepsilon \right) \times \tilde{\mathbb{P}} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2 \right\}, \end{aligned}$$

where $m := 0 \wedge \min_{a \in \mathcal{Y}} \tilde{\beta}(a, a)$. Therefore, by (3.3), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in O \right\} \\ & \geq -\langle \tilde{f}_\varpi, \varpi \rangle - \frac{1}{2} \langle \tilde{g}_\varpi, \varpi \rangle - \frac{1}{2} \langle \tilde{\beta}_\varpi, \varpi \otimes \varpi \rangle - \varepsilon + \liminf_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2 \right\}. \end{aligned}$$

The result follows once we prove that

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in B_\varpi^1 \times B_\varpi^2 \right\} = 0. \quad (3.8)$$

We use the upper bound (but now with the law \mathbb{P} replaced by $\tilde{\mathbb{P}}$) to prove (3.8). Then we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) \in (B_\varpi^1 \times B_\varpi^2)^c \right\} \leq - \inf_{(\tilde{\omega}, \tilde{\varpi}) \in \tilde{F}} \tilde{I}(\tilde{\omega}, \tilde{\varpi}),$$

where $\tilde{F} = (B_\varpi^1 \times B_\varpi^2)^c$ and $\tilde{I}(\tilde{\omega}, \tilde{\varpi}) := H(\tilde{\omega} \parallel \varpi) + \frac{1}{2} \mathfrak{H}_C^d(\tilde{\omega} \parallel \tilde{\varpi})$. It therefore suffices to show that the infimum is positive. Suppose for contradiction that there exists a sequence $(\tilde{\omega}_n, \tilde{\varpi}_n) \in \tilde{F}$ with $\tilde{I}(\tilde{\omega}_n, \tilde{\varpi}_n) \downarrow 0$. Then, because \tilde{I} is a good rate function and its level sets are compact, and by lower semi-continuity of the mapping $(\tilde{\omega}, \tilde{\varpi}) \mapsto \tilde{I}(\tilde{\omega}, \tilde{\varpi})$, we can construct a limit point $(\tilde{\omega}, \tilde{\varpi}) \in \tilde{F}$ with $\tilde{I}(\tilde{\omega}, \tilde{\varpi}) = 0$. By Lemma 3.4 this implies $H(\tilde{\omega} \parallel \varpi) = 0$ and $\mathfrak{H}_C(\tilde{\varpi} \parallel \tilde{\omega}) = 0$, hence $\tilde{\omega} = \varpi$, and $\tilde{\varpi} = \tilde{C}\tilde{\omega} \otimes \tilde{\omega} = \varpi$ contradicting $(\tilde{\omega}, \tilde{\varpi}) \in \tilde{F}$. \blacksquare

4. PROOF OF THEOREM 2.1

For any $n \in \mathbb{N}$ we define

$$\begin{aligned} \mathcal{P}_n(\mathcal{W}) &:= \left\{ \omega \in \mathcal{P}(\mathcal{W}) : n\omega(a) \in \mathbb{N} \text{ for all } a \in \mathcal{W} \right\}, \\ \tilde{\mathcal{P}}_n(\mathcal{W} \times \mathcal{W}) &:= \left\{ \varpi \in \tilde{\mathcal{P}}_*(\mathcal{W} \times \mathcal{W}) : \frac{n}{1 + \mathbb{1}_{\{a=b\}}} \varpi(a, b) \in \mathbb{N} \text{ for all } a, b \in \mathcal{W} \right\}. \end{aligned}$$

We denote by $\Theta_n := \mathcal{P}_n(\mathcal{W}) \times \tilde{\mathcal{P}}_n(\mathcal{W} \times \mathcal{W})$ and $\Theta := \mathcal{P}(\mathcal{W}) \times \tilde{\mathcal{P}}_*(\mathcal{W} \times \mathcal{W})$. With

$$\begin{aligned} P_{(\omega_n, \varpi_n)}^{(n)}(\mu_n) &:= \mathbb{P} \left\{ \mathcal{M}_X = \mu_n \mid \mathcal{H}(\mathcal{M}_X) = (\omega_n, \varpi_n) \right\}, \\ P^{(n)}(\omega_n, \varpi_n) &:= \mathbb{P} \left\{ (\mathcal{L}_X^1, \mathcal{L}_X^2) = (\omega_n, \varpi_n) \right\} \end{aligned}$$

the joint distribution of $\mathcal{L}_X^1, \mathcal{L}_X^2$ and \mathcal{M}_X is the mixture of $P_{(\omega_n, \varpi_n)}^{(n)}$ with $P^{(n)}(\omega_n, \varpi_n)$ defined as

$$d\tilde{P}^n(\omega_n, \varpi_n, \mu_n) := dP_{(\omega_n, \varpi_n)}^{(n)}(\mu_n) dP^{(n)}(\omega_n, \varpi_n). \quad (4.1)$$

(Biggins, Theorem 5(b), 2004) gives criteria for the validity of large deviation principles for the mixtures and for the goodness of the rate function if individual large deviation principles are known. The following three lemmas ensure validity of these conditions.

Lemma 4.1 (Doku-Amponsah, 2014b). *The family of measures $(P^n : n \in \mathbb{N})$ is exponentially tight on Θ*

Lemma 4.2 (Doku-Amponsah & Moerters, 2010). *The family of measures $(\tilde{P}^n : n \in \mathbb{N})$ is exponentially tight on $\Theta \times \mathcal{P}(\mathcal{W} \times \mathbb{N})$.*

Define the function

$$\tilde{J}: \Theta \times \mathcal{P}(\mathcal{W} \times \mathbb{N}) \rightarrow [0, \infty], \quad \tilde{J}((\omega, \varpi), \mu) = \tilde{J}_{(\omega, \varpi)}(\mu),$$

where

$$\tilde{J}_{((\varpi, \omega))}(\mu) = \begin{cases} H(\mu \| Q_{poi}) & \text{if } (\omega, \mu) \text{ is consistent and } \mu_1 = \omega_2 \\ \infty & \text{otherwise.} \end{cases} \quad (4.2)$$

Lemma 4.3 (Doku-Amponsah & Moerters, 2010). *\tilde{J} is lower semi-continuous.*

By (Biggins, Theorem 5(b), 2004) the two previous lemmas and the large deviation principles we have established in (Doku-Amponsah, Theorem 2.1, 2014b) and (Doku-Amponsah, Theorem 2.1, 2014c) ensure that under (\tilde{P}^n) the random variables $(\omega_n, \varpi_n, \mu_n)$ satisfy a large deviation principle on $\mathcal{P}(\mathcal{W}) \times \tilde{\mathcal{P}}_*(\mathcal{W} \times \mathcal{W}) \times \mathcal{P}(\mathcal{W} \times \mathbb{N})$ with good rate function

$$\hat{J}(\omega, \varpi, \mu) = \begin{cases} H(\omega \| \nu) + \frac{1}{2} \mathfrak{H}_C^d(\varpi \| \omega) + H(\mu \| Q_{poi}), & \text{if } (\varpi, \mu) \text{ is consistent and } \mu_1 = \varpi_2, \\ \infty, & \text{otherwise.} \end{cases}$$

By projection onto the last two components we obtain the large deviation principle as stated in Theorem 2.1 from the contraction principle, see e.g. (Dembo et al., 1998, Theorem 4.2.1).

5. PROOF OF COROLLARY 2.2, COROLLARY 2.3, AND COROLLARY 2.5

We derive the theorems from Theorem 2.1 by applying the contraction principle, see e.g. (Dembo & Zeitouni, Theorem 4.2.1, 1998). In fact Theorem 2.1 and the contraction principle imply a large deviation principle for D . It just remains to simplify the rate functions.

5.1 Proof of Theorem 2.2. Note that, in the case of an uncoloured RGG graphs, the function C degenerates to a constant c , $L^2 = |E|/n \in [0, \infty)$ and $M = D \in \mathcal{P}(\mathbb{N} \cup \{0\})$. Theorem 2.1 and the contraction principle imply a large deviation principle for D with good rate function

$$\eta_1(\delta) = \inf \{J(x, \delta) : x \geq 0\} = \inf \left\{ H(\delta \| q_x) + \frac{1}{2}x \log x - \frac{1}{2}x \log \rho(d)c + \frac{1}{2}\rho(d)c - \frac{1}{2}x : \langle \delta \rangle \leq x \right\},$$

which is to be understood as infinity if $\langle d \rangle$ is infinite. We denote by $\eta^x(\delta)$ the expression inside the infimum. For any $\varepsilon > 0$, we have

$$\eta^{\langle \delta \rangle + \varepsilon}(\delta) - \eta^{\langle \delta \rangle}(\delta) = \frac{\varepsilon}{2} + \frac{\langle \delta \rangle - \varepsilon}{2} \log \frac{\langle \delta \rangle}{\langle \delta \rangle + \varepsilon} + \frac{\varepsilon}{2} \log \frac{\langle \delta \rangle}{\rho(d)c} \geq \frac{\varepsilon}{2} + \frac{\langle \delta \rangle - \varepsilon}{2} \left(\frac{-\varepsilon}{\langle \delta \rangle} \right) + \frac{\varepsilon}{2} \log \frac{\langle \delta \rangle}{\rho(d)c} > 0,$$

so that the minimum is attained at $x = \langle \delta \rangle$.

5.2 Proof of Corollary 2.3. Corollary 2.3 follows from Theorem 2.2 and the contraction principle applied to the continuous linear map $G: \mathcal{P}(\mathbb{N} \cup \{0\}) \rightarrow [0, 1]$ defined by $G(\delta) = \delta(0)$. Thus, Theorem 2.2 implies the large deviation principle for $G(D) = W$ with the good rate function $\xi_1(y) = \inf \{ \eta_1(\delta) : \delta(0) = y, \langle \delta \rangle < \infty \}$. We recall the definition of η^x and observe that $\xi_2(y)$ can be expressed as

$$\xi_1(y) = \inf_{b \geq 0} \inf_{\substack{d \in \mathcal{P}(\mathbb{N} \cup \{0\}) \\ \delta(0)=y, \rho(d)c\langle \delta \rangle=b^2}} \left\{ \frac{1}{2}c + y \log y + \frac{b^2}{2\rho(d)c} + \sum_{k=1}^{\infty} \delta(k) \log \frac{\delta(k)}{q_b(k)} - b(1-y) \right\}.$$

Now, using Jensen's inequality, we have that

$$\sum_{k=1}^{\infty} \delta(k) \log \frac{\delta(k)}{q_b(k)} \geq (1-y) \log \frac{(1-y)}{(1-e^{-b})}, \quad (5.1)$$

with equality if $\delta(k) = \frac{(1-y)}{(1-e^{-b})} q_b(k)$, for all $k \in \mathbb{N}$. Therefore, we have the inequality

$$\inf \{ \eta(\delta) : \delta(0) = y, \langle \delta \rangle < \infty \} \geq \inf \left\{ \frac{1}{2}c + y \log y + \frac{b^2}{2\rho(d)c} + (1-y) \log \frac{(1-y)}{(1-e^{-b})} - b(1-y) : b \geq 0 \right\}.$$

Let $y \in [0, 1]$. Then, the equation $a(1-e^{-a}) = \rho(d)c(1-y)$ has a unique positive solution. Elementary calculus shows that the global minimum of $b \mapsto \frac{1}{2}\rho(d)c + y \log y + \frac{b^2}{2\rho(d)c} + (1-y) \log \frac{(1-y)}{(1-e^{-b})} - b(1-y)$ on $(0, \infty)$ is attained at the value $b = a$, where a is the positive solution of our equation. We obtain the form of ξ in Corollary 2.3 by observing that

$$\frac{a(y)^2 + (\rho(d)c)^2 - 2\rho(d)ca(y)(1-y)}{2\rho(d)c} = \frac{\rho(d)cy}{2}(2-y) + \frac{1}{2\rho(d)c}(a(y) - \rho(d)c(1-y))^2.$$

5.3 Proof of Corollary 2.5. We define the continuous linear map $W : \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y}^2) \rightarrow [0, \infty)$ by $W(\varpi, \varpi) = \frac{1}{2}\|\varpi\|$, and infer from Theorem 2.4 and the contraction principle that $W(\mathcal{L}_X^1, \mathcal{L}_X^2) = |E|/n$ satisfies a large deviation principle in $[0, \infty)$ with the good rate function

$$\zeta(x) = \inf \{ I(\varpi, \varpi) : W(\varpi, \varpi) = x \}.$$

To obtain the form of the rate in the corollary, the infimum is reformulated as unconstrained optimization problem (by normalising ϖ)

$$\inf_{\substack{\varpi \in \mathcal{W}_*(\mathcal{Y}^2) \\ \varpi \in \mathcal{W}(\mathcal{Y})}} \left\{ H(\varpi \| \nu) + xH(\varpi \| \Delta(d)C\varpi \otimes \varpi) + x \log 2x + \frac{\Delta(d)}{2} \|C\varpi \otimes \varpi\| - x \right\}. \quad (5.2)$$

By Jensen's inequality $H(\varpi \| \Delta(d)C\varpi \otimes \varpi) \geq -\log \|\Delta(d)C\varpi \otimes \varpi\|$, with equality if $\varpi = \frac{C\varpi \otimes \varpi}{\|C\varpi \otimes \varpi\|}$, and hence, by symmetry of C we have

$$\begin{aligned} \min_{\varpi \in \mathcal{W}_*(\mathcal{Y}^2)} & \left\{ H(\varpi \| \nu) + xH(\varpi \| \Delta(d)C\varpi \otimes \varpi) + x \log 2x + \frac{\Delta(d)}{2} \|C\varpi \otimes \varpi\| - x \right\} \\ & = H(\varpi \| \nu) - x \log \|\Delta(d)C\varpi \otimes \varpi\| + x \log 2x + \frac{\Delta(d)}{2} \|C\varpi \otimes \varpi\| - x. \end{aligned}$$

The form given in Corollary 2.5 follows by defining

$$y = \Delta(d) \sum_{a,b \in \mathcal{Y}} C(a,b) \varpi(a) \varpi(b).$$

REFERENCES

- J.D. BIGGINS.(2004) Large deviations for mixtures. *El. Comm. Probab.* **9** 60–71 (2004).
J.D. BIGGINS and D.B. PENMAN.(2009) Large deviations in randomly coloured random graphs. *Electron. Comm. Probab.* **14** 290-301 (2009).
S. BOUCHERON, O. BOUSQUET and G. LUGOSI.(2004) Concentration inequalities. In: *Advanced lectures in machine learning*, Eds: O. Bousquet, U. v. Luxburg and G. R  tsch. Springer (2004) 208-240.
S. BOUCHERON, F. GAMBOA and C. LEONARD.(2002) Bins and balls: Large deviations of the empirical occupancy process. *Ann. Appl. Probab.* **12** 607-636 (2002).
C. CANNINGS and D.B. PENMAN.(2003) Models of random graphs and their applications. In: *Handbook of Statistics 21. Stochastic Processes: Modelling and Simulation*. Eds: D.N. Shanbhag and C.R. Rao. Elsevier (2003) 51-91.

- A. DEMBO, P. MÖRTERS and S. SHEFFIELD.(2005) Large deviations of Markov chains indexed by random trees. *Ann. Inst. Henri Poincaré: Probab. et Stat.* **41** 971-996 (2005).
- A. DEMBO and O. ZEITOUNI.(1998) *Large deviations techniques and applications*. Springer, New York, (1998).
- K. DOKU-AMPONSAH.(2006) *Large deviations and basic information theory for hierarchical and networked data structures*. PhD Thesis, Bath (2006).
- K. DOKU-AMPONSAH.(2012) Asymptotic equipartition principles for simple hierarchical and networked data structures. *Esaim: Probability and Statistics* **16**, 114-138 (2012).
- K. DOKU-AMPONSAH.(2014[a]) Exponential approximation, method of types for empirical neighbourhood measures of random graphs by random allocation. <http://arxiv.org/pdf/1212.4281.pdf>
- K. DOKU-AMPONSAH.(2014[b]) Large deviations for number of edges of near intermediate random geometric graphs. Unpublished Manuscript (2012).
- K. DOKU-AMPONSAH.(2014[c]) Large deviation result for the empirical locality measure of typed random geometric graphs. Unpublished Manuscript (2012).
- K. DOKU-AMPONSAH and P. MÖRTERS.(2010) Large Deviation Principles for empirical Measures of Coloured random graphs. *The annals of Applied Probab.* Vol. 20, No. 6, 1989-2021. DOI:10.1214/09-AAP47.
- T. MILENKOVIC. *Graph-theoretical approaches for studying biological networks*. Unplished manuscript, Department of Computer Science University of California, Irvine
- C.J.H. MCDIARMID and T. MÜLLER.(2005) Colouring random geometric graphs DMTCS proc. AE, 2005, 1-4.
- S. MUKHERJEE(2014) Large deviation for the empirical degree distribution of an Erdos-Renyi graph.
- N. O'CONNELL.(1998) Some large deviation results for sparse random graphs. *Probab. Theory Relat. Fields* **110** 277-285 (1998).
- D.B. PENMAN.(1998) *Random graphs with correlation structure*. PhD Thesis, Sheffield 1998.
- M.D. PENROSE.(2003) *Random geometric Graphs*. Oxford University press. Oxford (2003).